

BỘ GIÁO DỤC VÀ ĐÀO TẠO BỘ KHOA HỌC VÀ CÔNG NGHỆ
VIỆN NĂNG LƯỢNG NGUYÊN TỬ VIỆT NAM



HỒ QUANG TUẤN

**NGHIÊN CỨU ỨNG DỤNG HỌC MÁY TRONG XỬ LÝ ẢNH PET/CT
LÒNG NGỰC HỖ TRỢ CÔNG TÁC CHẨN ĐOÁN UNG THƯ PHỔI**

TÓM TẮT LUẬN ÁN TIẾN SĨ VẬT LÝ

Chuyên ngành: Vật lý nguyên tử và hạt nhân

Mã số: 9.44.01.06

Người hướng dẫn khoa học:

PGS.TS. Nguyễn Văn Thái

TS. Lê Xuân Chung

Hà Nội – 2026

Công trình được hoàn thành tại:

Viện khoa học kỹ thuật hạt nhân, Viện Năng lượng nguyên tử Việt Nam

Người hướng dẫn khoa học:

1. HDC: PGS.TS. Nguyễn Văn Thái - Đại học Bách khoa Hà nội

2. HDP: TS. Lê Xuân Chung - Viện năng lượng nguyên tử Việt Nam

Phản biện 1:

Phản biện 2:

Phản biện 3:

Phản biện độc lập 1:

Phản biện độc lập 2:

Luận án sẽ được bảo vệ trước Hội đồng chấm luận án cấp Viện họp tại Trung tâm đào tạo hạt nhân, Viện Năng lượng nguyên tử Việt Nam, vào lúc.....ngày.....tháng.....năm 20...

Có thể tìm hiểu luận án tại thư viện:

1. Thư viện Tổng hợp Quốc gia

2. Thư viện Trung tâm đào tạo - Viện Năng lượng nguyên tử Việt Nam

MỞ ĐẦU

Chẩn đoán hình ảnh hiện đại đóng vai trò quyết định trong việc cung cấp các bằng chứng định lượng phục vụ xây dựng phác đồ và đánh giá đáp ứng điều trị ung thư. Sự phát triển của các hệ thống ghi hình y tế thế hệ mới đã tối ưu hóa đáng kể độ phân giải giải phẫu và chức năng, song cũng làm phát sinh thách thức về mặt xử lý "dữ liệu lớn" (Big Data). Điển hình là trong tầm soát và quản lý ung thư phổi, kỹ thuật chụp ảnh cắt lớp phát xạ positron kết hợp vi tính (PET/CT) đã trở thành tiêu chuẩn vàng nhờ năng lực hợp nhất thông tin cấu trúc (ảnh CT) và thông tin chuyển hóa sinh học (ảnh PET). Mặc dù các hệ thống PET/CT đã tích hợp các phần mềm hỗ trợ xử lý dữ liệu thông minh, việc khai thác hiệu quả khối dữ liệu thể tích 3D tích hợp lên tới hàng ngàn lát cắt này vẫn đang gặp phải những nút thắt lớn do quy trình phân tích lâm sàng truyền thống chủ yếu dựa vào đánh giá định tính của con người.

Nhằm khắc phục các hạn chế này, xu hướng ứng dụng học máy (Machine Learning) và học sâu (Deep Learning) trong xử lý ảnh PET/CT đang trở thành một hướng nghiên cứu mũi nhọn, đặc biệt là xu hướng Radiomics. Cụ thể, dữ liệu ảnh đa phương thức PET/CT đóng vai trò là nền tảng cốt lõi định hình độ chính xác, tính khả thi và độ an toàn sinh học của các mô hình học sâu. Nếu thiếu nguồn dữ liệu PET/CT được chuẩn hóa cấu trúc và gán nhãn chuẩn mực (ground truth), các mô hình AI chẩn đoán sẽ không thể vượt qua ranh giới thử nghiệm phòng thí nghiệm để đạt đủ điều kiện phê duyệt ứng dụng trong thực tiễn lâm sàng. Phần lớn các bộ dữ liệu huấn luyện hiện nay đều có nguồn gốc từ các quần thể phương Tây, đẩy lên những lo ngại sâu sắc về tính bền vững của thuật toán khi triển khai trên các bệnh nhân châu Á — những người vốn sở hữu các đặc điểm di truyền, đặc tính bệnh lý và phân phối giá trị hấp thụ chuẩn hóa (SUV) hoàn toàn khác biệt.

Trong bối cảnh đó, luận án với đề tài **"Nghiên cứu ứng dụng học máy trong xử lý ảnh PET/CT lồng ngực hỗ trợ công tác chẩn đoán bệnh ung thư phổi"** được thực hiện tập trung vào ba mục tiêu chính sau:

- (1) Thiết lập cơ sở dữ liệu ảnh chụp PET/CT người Việt bản địa quy mô lớn
- (2) Xây dựng công cụ đọc và tiền xử lý ảnh nhằm trích xuất dữ liệu ảnh DICOM phục vụ huấn luyện mô hình học sâu
- (3) Phát triển nền tảng học sâu tự động phân đoạn ảnh chụp PET/CT hỗ trợ công tác chẩn đoán ung thư phổi

Ngoài phần mở đầu, kết luận và tài liệu tham khảo, nội dung chính của luận án được trình bày trong 3 chương: **Chương I** trình bày nghiên cứu tổng quan, **Chương II** trình bày phương pháp nghiên cứu, **Chương III** trình bày kết quả khảo sát đánh giá và bàn luận.

CHƯƠNG 1. NGHIÊN CỨU TỔNG QUAN

1.1. Vai trò của ảnh chụp PET/CT trong chẩn đoán ung thư phổi

So với SPECT, kỹ thuật PET sở hữu lợi thế vượt trội nhờ khả năng khai thác đa dạng các đồng vị phóng xạ phát positron. Khi được gắn kết với các phân tử sinh học hoặc dược phẩm, các đồng vị này đóng vai trò như những chất đánh dấu phân tử (molecular tracers). Do các tế bào ung thư thường biểu hiện quá mức các kênh vận chuyển GLUT và có hoạt tính enzyme Hexokinase cao hơn hẳn tế bào bình thường để phục vụ nhu cầu năng lượng lớn, chúng sẽ tích tụ một lượng dược chất phóng xạ lớn. Sự chênh lệch về nồng độ phóng xạ này cho phép máy PET phát hiện và tái tạo thành những hình ảnh sáng tương phản, giúp các bác sĩ xác định chính xác vị trí và mức độ hoạt động của các tổn thương bệnh lý trên cơ thể người bệnh. Đánh giá định tính trong ghi hình ^{18}F -FDG PET/CT tập trung vào việc phân tích trực quan mức độ hấp thụ dược chất phóng xạ dựa trên **thang điểm Herder 4 mức độ**, thay vì dựa vào các con số thuần túy.

Chỉ số SUV (Standardized Uptake Value): Chỉ số hấp thụ chuẩn hóa (Standardized Uptake Value - SUV) là một đại lượng tính toán được sử dụng trong ghi hình với ^{18}F -FDG nhằm nỗ lực chuẩn hóa mức độ hấp thụ của dược chất phóng xạ này dựa trên liều lượng tiêm vào, sự phân rã phóng xạ tính từ thời điểm tiêm và khối lượng cơ thể của bệnh nhân

Thể tích khối u chuyển hóa (MTV - Metabolic Tumor Volume): Định nghĩa của MTV là thể tích nằm trong một vùng quan tâm (ROI) do người dùng hoặc thuật toán xác định nhằm phân tách phần khối u có hoạt động chuyển hóa. Để xác định ranh giới của ROI này, các phương pháp dựa trên ngưỡng (threshold-based) hoặc dựa trên thuật toán (algorithm-based) đã được đề xuất và đánh giá. Rất nhiều thuật toán phân đoạn PET đã được phát triển và ứng dụng cho hình ảnh FDG PET. Việc phân nhóm các thuật toán này thành các lớp riêng biệt là một thách thức do sự đa dạng cực kỳ lớn của chúng.

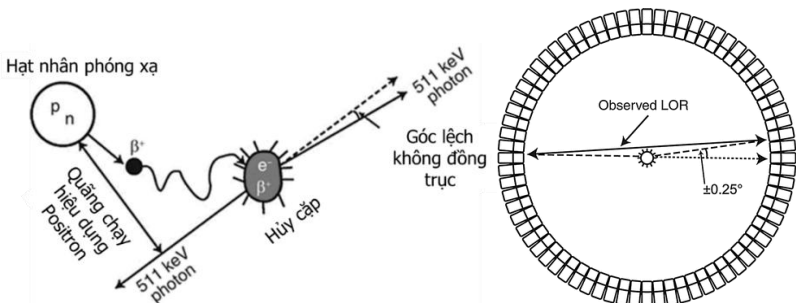
Tổng lượng glycolysis tổn thương (TLG - Total Lesion Glycolysis): là tích số giữa SUV trung bình (mean SUV) và thể tích khối u chuyển hóa (MTV) được sử dụng nhằm đánh giá đáp ứng điều trị. Chỉ số này là sự kết hợp giữa thông tin về thể tích và thông tin về chuyển hóa trên hình ảnh FDG PET.

Việc sử dụng PET-FDG trong ung thư học để xác định giai đoạn và chẩn đoán tái phát khối u đang là một trong những lĩnh vực phát triển nhanh chóng nhất của y học hạt nhân nói riêng và chẩn đoán hình ảnh nói chung. Ứng dụng của PET-FDG trong chẩn đoán phân biệt nốt đơn độc ở phổi, cũng như xác định giai đoạn và phát hiện tái phát ung thư phổi, ung thư đại trực tràng đóng vai trò đặc biệt quan trọng trong ung thư học lâm sàng. Giá trị của kỹ thuật PET trong việc phát hiện sự hiện diện của tế bào ung thư trong những tình huống này là kết quả của sự cải thiện vượt bậc về độ nhạy đối với các khối u thể tích nhỏ, dựa trên những tiến bộ gần đây của thiết bị ghi hình PET.

1.2. Đặc trưng dữ liệu ảnh PET/CT và các tác vụ xử lý cốt lõi

Chất lượng ảnh trong hệ thống chụp ảnh PET/CT phụ thuộc vào rất nhiều yếu tố vật lý phức tạp. Do đây là hệ thống tích hợp phương thức ảnh CT và ảnh PET nên các yếu tố ảnh hưởng sẽ chia thành ba nhóm chính: bản chất vật lý của phân rã phóng xạ, các hiện tượng tương tác bức xạ và các yếu tố kỹ thuật của hệ thống.

Bản chất của PET là ghi nhận hai photon hủy cặp (511 keV) phát ra theo hai hướng ngược chiều nhau. Tuy nhiên, trước khi quá trình hủy cặp xảy ra, chuyển động của hạt Positron gây ảnh hưởng đến độ phân giải không gian theo hai cách (Hình 1.1):



Hình 1.1: Ảnh hưởng vị trí và góc lệch không đồng trục tới độ phân giải

Sau khi phát ra từ hạt nhân, positron di chuyển một quãng đường ngắn trong mô để tiêu hao động năng trước khi hủy cặp với electron. Vị trí hủy cặp (nơi phát ra photon) lệch một khoảng so với vị trí thực tế, tùy vào hạt nhân phóng xạ. Ví dụ như ^{18}F có năng lượng tối đa thấp (0.63 MeV), quãng chạy trung bình trong nước chỉ khoảng 0.6 mm, cho độ phân giải tốt hơn so với ^{68}Ga (1.90 MeV, quãng chạy ~ 2.9 mm). Ngoài ra, theo lý thuyết, hai photon phát ra theo hai hướng ngược nhau, tạo góc 180° . Thực tế, do positron vẫn còn một động năng cực nhỏ (mô-men động lượng dư), góc hủy cặp bị lệch khoảng 0.5° . Với đường kính vòng detector lớn (ví dụ 80 - 90 cm), độ lệch này có thể làm nhòe ảnh khoảng 1.5 – 2 mm.

Hệ thống PET ghi nhận ảnh dựa trên các sự kiện trùng phùng đồng thời (True Coincidences). Tuy nhiên, các tương tác vật lý của photon 511 keV với mô cơ thể sinh ra các loại nhiễu nghiêm trọng do Hiện tượng tán xạ (Scattered Coincidences), sự trùng phùng ngẫu nhiên (Random Coincidences), sự đâm xuyên detector (Parallax Effect / Depth of Interaction – DOI). Ngoài ra còn phải kể đến các hiệu ứng ảnh hưởng khác như sự suy giảm chùm tia (Attenuation), các yếu tố phần cứng của hệ thống Detector và hiệu ứng thể tích bán phần (Partial Volume Effect - PVE)

Quy trình xử lý ảnh PET/CT (chuyển hóa - giải phẫu) bao gồm một chuỗi các tác vụ tính toán phức tạp. Các tác vụ này trải dài từ giai đoạn tiền xử lý (nhận dạng, chuẩn hóa), tái cấu trúc cho đến phân tích định lượng nâng cao phục vụ lâm sàng và nghiên cứu.

Tác vụ 1. Đăng ký ảnh và hợp nhất ảnh (Image Registration & Fusion)

Tác vụ 2. Tiền xử lý và Hiệu chỉnh ảnh (Pre-processing & Image Correction)

Tác vụ 3. Tái cấu tạo ảnh (Image Reconstruction)

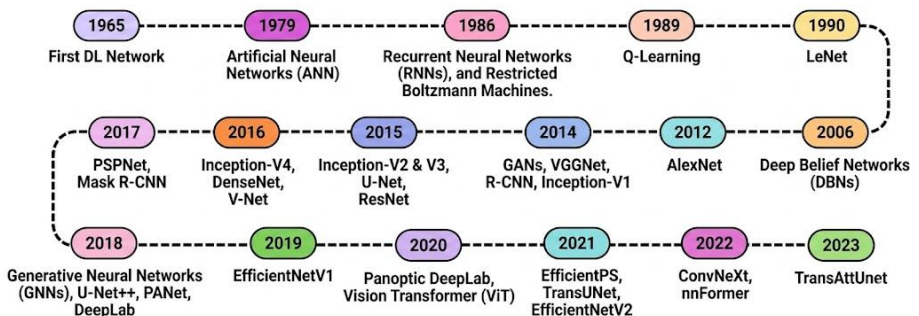
Tác vụ 4. Hiệu chỉnh chuyển động (Motion Correction / Gating)

Tác vụ 5. Phân tích định lượng và Trích xuất (Quantification & Radiomics)

1.3. Ứng dụng học máy trong xử lý ảnh PET/CT

Mạng tích chập (Convolutional Neural Network - CNN) là một trong những kiến trúc học sâu (Deep Learning) tiên tiến và phổ biến nhất hiện nay. Nó được thiết kế đặc biệt để xử lý các dữ liệu có cấu trúc lưới hai chiều, điển hình là **hình ảnh**. Thay vì xem một bức ảnh là một dải phẳng các điểm ảnh độc lập, CNN có khả năng hiểu được mối quan hệ không gian giữa các pixel cạnh nhau, từ đó tự động nhận diện các đặc trưng từ đơn giản (cạnh, góc) đến phức tạp (hình dáng, vật thể). Trong một tầng tích chập, chúng ta huấn luyện nhiều bộ lọc khác nhau để trích xuất các bản đồ đặc trưng đa dạng từ hình ảnh đầu vào. Khi chúng ta xếp chồng nhiều tầng tích chập liên tiếp nhau, kết hợp cùng các hàm phi tuyến, chúng ta sẽ tạo nên một **mạng thần kinh tích chập (CNN)**. Như vậy, mỗi tầng tích chập thực hiện đồng thời hai nhiệm vụ: 1) **Lọc không gian (Spatial filtering)**: Thông qua phép toán tích chập giữa hình ảnh và các hạt nhân (kernels) và 2) **Kết hợp đa kênh (Combining multiple input channels)**: Tổng hợp các kênh đầu vào để tạo ra một tập hợp các kênh đầu ra mới. Có đến 90% các nghiên cứu về CNN tập trung vào việc sửa đổi hoặc cải thiện chính hai khía cạnh này.

Hình 1.2 minh họa lộ trình phát triển của các mô hình học sâu (DL) trong phân đoạn hình ảnh y tế, phản ánh những cải tiến liên tục về mặt chất lượng cũng như tính cấp thiết của công nghệ này.



Hình 1.2: Các mô hình học sâu trong phân đoạn hình ảnh

Theo thời gian, những bước tiến này đã nâng cao đáng kể năng lực của các mô hình học sâu. Khả năng xử lý từ đầu đến cuối (end-to-end), xử lý không gian 3D và học chuyển giao (transfer learning) đã mang lại kết quả phân đoạn chính xác vượt trội; trong khi các phương pháp như giám sát yếu, tự giám sát và mô hình học máy

kết hợp (ensemble modeling) đã giải quyết hiệu quả các thách thức về sự khan hiếm dữ liệu có nhãn cũng như sự mất cân bằng dữ liệu giữa các lớp. Khả năng triển khai trong thời gian thực, học liên tục và tính thích ứng cao đã gia tăng giá trị ứng dụng lâm sàng, nơi các mô hình như U-NET, U-NET++ và FCN (Mạng tích chập hoàn toàn) đã trở thành những trụ cột trong lĩnh vực phân đoạn hình ảnh y tế. Sự phát triển không ngừng này phản ánh tầm quan trọng và khả năng ứng dụng ngày càng rộng rãi của các kỹ thuật trên trong lĩnh vực y học. Phân đoạn hình ảnh đã trải qua một hành trình tiến hóa ngoạn mục, từ những phương pháp luận Máy học (ML) truyền thống cho đến trạng thái phát triển bùng nổ chưa từng có trong kỷ nguyên Học sâu hiện nay. Theo số liệu nghiên cứu tổng quan, các kiến trúc được ưa chuộng nhất trong các nghiên cứu lần lượt là U-net (36 nghiên cứu), ResNet (20 nghiên cứu) và GAN (16 nghiên cứu). Đáng chú ý, có 12 bài báo đã ứng dụng kỹ thuật học chuyển giao (transfer learning) và 9 bài báo đưa đặc trưng Radiomics vào quy trình nghiên cứu.

1.4. Đối tượng và mục tiêu nghiên cứu của luận án

Vấn đề mất cân bằng dữ liệu là một thách thức lớn trong các nghiên cứu radiomics phục vụ chẩn đoán ung thư phổi lành tính và ác tính, dễ dẫn đến kết quả chẩn đoán bị thiên lệch (biased). Mặc dù hiện có một vài bộ dữ liệu PET/CT công khai quy mô lớn (ví dụ: AutoPET, LUNG-PET-CT-Dx), song chúng thường thiếu sự xác thực toàn diện về mặt mô bệnh học, hoặc chưa được tối ưu hóa để đánh giá hiện tượng dịch chuyển miền dữ liệu (domain shift) giữa các sắc tộc đa dạng. Phần lớn các bộ dữ liệu huấn luyện hiện nay đều có nguồn gốc từ các quần thể phương Tây, dấy lên những lo ngại sâu sắc về tính bền vững (robustness) của thuật toán khi triển khai trên các bệnh nhân châu Á — những người vốn sở hữu các đặc điểm di truyền, đặc tính bệnh lý và phân phối giá trị hấp thụ chuẩn hóa (SUV) hoàn toàn khác biệt. Hơn nữa, trong ảnh PET/CT, khối u nguyên phát và các hạch vùng hoặc di căn xa có mối quan hệ bệnh lý chặt chẽ. CNN thuần túy không có khả năng mô hình hóa tường minh mối quan hệ phụ thuộc tầm xa (long-range dependencies) này, dẫn đến việc phân đoạn rời rạc và thiếu tính nhất quán mang tính hệ thống của ca bệnh. Hơn nữa, CNN học các mẫu kết cấu (texture) cục bộ rất mạnh mẽ. Ảnh PET vốn có bản chất là độ phân giải thấp, tỷ lệ tín hiệu trên nhiễu (SNR) kém và dễ bị ảnh hưởng bởi chuyển động hô hấp của bệnh nhân (gây nhòe ảnh). Trong khi đó, ảnh CT có thể bị nhiễu do các dị vật kim loại (răng giả, nẹp xương). Các bộ lọc tích chập cục bộ của CNN rất dễ bị đánh lừa bởi các vùng nhiễu này, dẫn đến kết quả phân đoạn xuất hiện nhiều vùng dương tính giả (false positives).

Nhằm đề xuất giải pháp phù hợp giải quyết những hạn chế nêu trên, nội dung nghiên cứu của luận án tập trung vào ba mục tiêu chính sau:

- (1) Thiết lập cơ sở dữ liệu ảnh chụp PET/CT người Việt bản địa quy mô lớn
- (2) Xây dựng công cụ đọc và tiền xử lý ảnh nhằm trích xuất dữ liệu ảnh DICOM phục vụ huấn luyện mô hình học sâu

- (3) Phát triển nền tảng học sâu tự động phân đoạn ảnh chụp PET/CT hỗ trợ công tác chẩn đoán ung thư phổi

CHƯƠNG 2. PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Xây dựng bộ dữ liệu người Việt

Mục tiêu chính của việc thu thập dữ liệu ảnh PET/CT là tạo ra một bộ dữ liệu đại diện, chất lượng cao, có nhãn chính xác để huấn luyện, hiệu chỉnh và đánh giá các mô hình trí tuệ nhân tạo (AI), đặc biệt trong chẩn đoán và hỗ trợ ra quyết định lâm sàng đối với ung thư phổi không tế bào nhỏ (NSCLC) và các bệnh lý liên quan. Dữ liệu sẽ được sử dụng để phát triển các mô hình như: Phân đoạn tổn thương (tumor segmentation), Phân loại tổn thương (classification: ác tính/lành tính), Dự đoán giai đoạn bệnh (staging), Ước tính khả năng sống còn hoặc đáp ứng điều trị. Phạm vi dữ liệu bao gồm: Ảnh PET/CT của bệnh nhân đã được chẩn đoán xác định (confirmed diagnosis), ảnh từ hệ thống máy PET/CT các bệnh viện tại các khu vực khác nhau có thể bao gồm cả hình ảnh của bệnh nhân không có bệnh lý phổi ác tính để làm mẫu đối chứng (negative control). Các nguồn dữ liệu được khai thác hợp pháp và đảm bảo tính đại diện: dữ liệu lưu trữ lịch sử của các Bệnh viện, có văn bản đồng ý tham gia nghiên cứu của các Bệnh viện.

Để đảm bảo tính ứng dụng và độ tin cậy của mô hình AI, dữ liệu PET/CT cần đáp ứng các đặc điểm sau: đa dạng về bệnh lý (bệnh nhân ung thư phổi không tế bào nhỏ giai đoạn khác nhau, có di căn xa hay di căn hạch, trường hợp nốt phổi đơn độc, ...), đa dạng về thiết bị (dữ liệu từ các đơn vị khác nhau), đa dạng dân số (phân bố hợp lý độ tuổi, giới tính, hút thuốc hay không hút thuốc, nhóm nguy cơ cao), độ phân giải ảnh PET không thấp hơn 128x128, ảnh CT từ 512x512 (bề dày lớp cắt $\leq 5\text{mm}$), không bị nhiễu mạnh, không có lỗi về trường nhìn hay cắt cụt hình ảnh. Để đảm bảo bảo mật các trường PatientName, PatientID, BirthDate được xóa hoặc mã hóa, dữ liệu được lưu trữ nội bộ, có phân quyền truy cập.

Các tiêu chí lựa chọn được xây dựng dựa trên hướng dẫn của các tổ chức như EANM, AJCC Cancer Staging Manual và Fleischner Society để đảm bảo tính chính xác và phù hợp với mục tiêu phân loại NSCLC hoặc không phải NSCLC. Theo kết quả giải phẫu bệnh cuối cùng, các trường hợp dữ liệu thu thập được chia thành nhóm dương tính và nhóm âm tính với ung thư không tế bào nhỏ. Nghiên cứu hồi cứu tuân theo một quy trình đã được Hội đồng Đánh giá Tổ chức tại Bệnh viện phê duyệt và không cần sự đồng ý của bệnh nhân. Đây là cũng một trong các tiêu chí chung để lựa chọn bộ dữ liệu đang được sử dụng trong các nghiên cứu hiện nay. Nghiên cứu sinh đã thực hiện việc thu thập bộ dữ liệu có sẵn, công khai về các nghiên cứu Chụp cắt lớp phát xạ Positron/Chụp cắt lớp điện toán (PET/CT) từ năm 2019 đến năm 2023. Tất cả dữ liệu ảnh đều được thực hiện trên máy quét PET/CT hiện đại của Bệnh viện **Ung Bướu Hà Nội, Bệnh viện K và Bệnh viện Ung Bướu TP. Hồ Chí Minh**. Việc thu thập dữ liệu PET/CT này đã được sự đồng ý bằng văn bản của 03 bệnh viện trên.

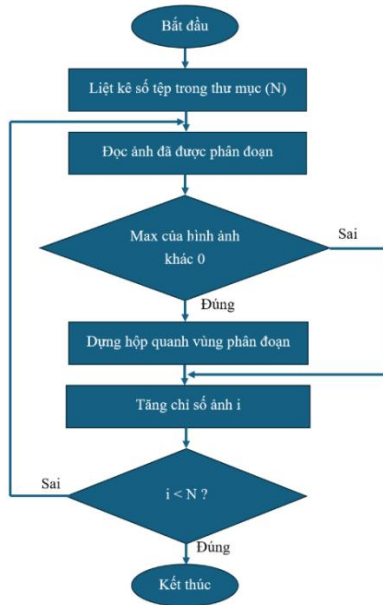
Các hệ thống chụp PET/CT của 3 bệnh viện đều có model là Discovery IQ của GE. Quy trình chụp dựa theo hướng dẫn của Bộ Y tế tại Quyết định 705/QĐ-BYT ngày 28/02/2014. Quá trình thu thập thông tin được thực hiện một cách có hệ thống, bao gồm hình ảnh PET/CT và dữ liệu lâm sàng. Hình ảnh PET/CT là thành phần cốt lõi của cơ sở dữ liệu, bao gồm dữ liệu từ hai phương thức hình ảnh bổ sung cho nhau. Để đánh giá chất lượng của bộ dữ liệu PET/CT phục vụ ứng dụng học máy trong chẩn đoán ung thư phổi, một số tiêu chí quan trọng nhằm đảm bảo dữ liệu phù hợp, đáng tin cậy và hỗ trợ hiệu quả cho việc huấn luyện mô hình như: **Độ chính xác và độ tin cậy của nhãn, Chất lượng hình ảnh, Tính đa dạng và đại diện, Kích thước bộ dữ liệu, Tính nhất quán và chuẩn hóa, Tính đầy đủ của thông tin bổ trợ, Tính pháp lý và đạo đức, Khả năng sử dụng cho học máy.**

2.2. Xây dựng công cụ đọc và tiền xử lý ảnh

Tiền xử lý ảnh PET/CT là một bước quan trọng trong quá trình phân tích và đánh giá hình ảnh y tế bao gồm các bước: ***Cân bằng độ sáng, Lọc, Đồng ký ảnh, Chuẩn hóa cường độ, Xử lý chuyển đổi, Loại bỏ ảnh giả.***

Các khối u có trong dữ liệu hình ảnh PET/CT của các bệnh nhân được bác sĩ đánh dấu thủ công bằng cách khoanh và tô kín diện tích khối u. Trong quá trình tô, các bác sĩ có thể lựa chọn gán thẻ cho từng vùng khoanh, phần mềm phân đoạn thủ công sẽ gán giá trị tương ứng cho các điểm ảnh mà bác sĩ đã gán thẻ. Trong luận án, các điểm ảnh của khối u có giá trị bằng 1, phần mềm đọc và hiển thị khối u sẽ đọc và xác định khu vực có chứa các điểm ảnh có kích thước bằng 1 và bao bọc quanh khối u bằng một hình chữ nhật. Việc xác định đường biên dạng chính xác của khối u là không cần thiết vì máy tính chỉ có thể lấy hình ảnh là hình vuông hoặc chữ nhật để đưa vào mô hình huấn luyện. Lưu đồ thuật toán của chương trình được thể hiện trên [Hình 2.1](#). Để xác định đường bao của khối u, chương trình sẽ thực hiện tính vi phân theo từng hàng, cột của hình ảnh phân đoạn. Kết quả, ta sẽ thu được các điểm mà tại đó cường độ thay đổi một cách đột ngột. Biên dạng của khối u có thể được xác định bằng cách lấy các điểm vi phân có vị trí liên kề. Với thuật toán này, ta sẽ xác định được rất nhiều đường biên dạng của khối u vì các khối u có thể được tô không kín hoặc khối u có hình dạng phân nhánh (kết quả trên một lát cắt ta sẽ nhìn thấy hai vùng được tô kín). Đối với các khối u không được tô kín, chương trình sẽ phát hiện được rất nhiều biên dạng nằm bên trong một biên dạng lớn. Để xử lý được trường hợp này, nghiên cứu sinh tiến hành loại bỏ các đường bao mà có thể tích nhỏ hơn 5 mm, do mục tiêu của luận án chỉ phát hiện các nốt có kích thước lớn cỡ centimet. Các đường bao quanh vùng diện tích lớn hơn 5 mm sẽ được tiếp tục xử lý dựa trên tọa độ của tâm đường bao, nếu tọa độ tâm của đường bao nằm bên trong một đường bao lớn hoặc không cách đường bao lớn quá 5 mm thì sẽ được coi là các điểm khoanh lỗi hoặc thuộc về đường bao lớn. Các đường bao sau khi được xử lý sẽ hợp nhất lại thành các đường bao thỏa mãn điều kiện của luận án. Đối với các khối u được khoanh thành nhiều phần kín, liên tục, chương

trình sẽ tiến hành xác định diện tích và tâm của khối u, nếu khoảng cách tâm của các đường bao khối u không lớn hơn kích thước của các đường bao đó, thì các đường bao sẽ được hợp nhất và coi là một khối u duy nhất (Hình 2.2).



Hình 0.1: Lưu đồ thuật toán chương trình xác định đường bao



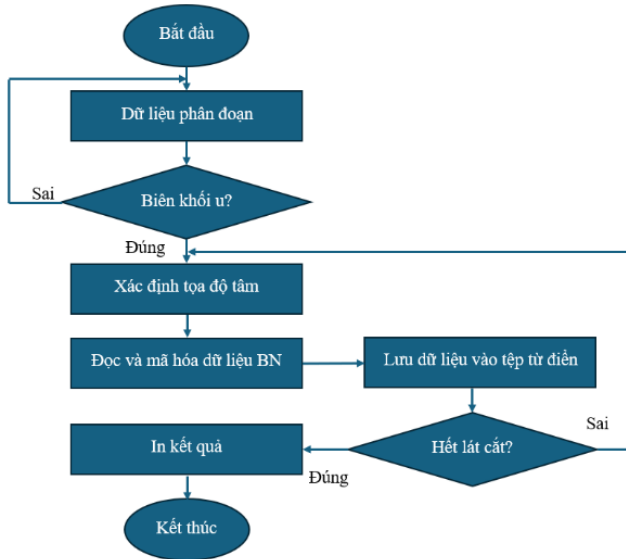
a)



b)

Hình 2.2. Hình ảnh khối u a) không được tô kín và b) không phân nhánh

Lưu đồ chương trình ghi thông tin kết quả phân đoạn thủ công được thể hiện như trong Hình 2.3.

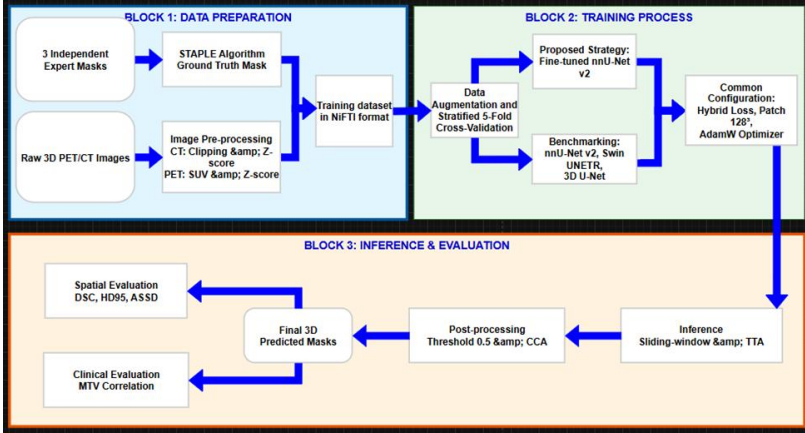


Hình 2.3. Lưu đồ chương trình ghi thông tin kết quả phân đoạn thủ công

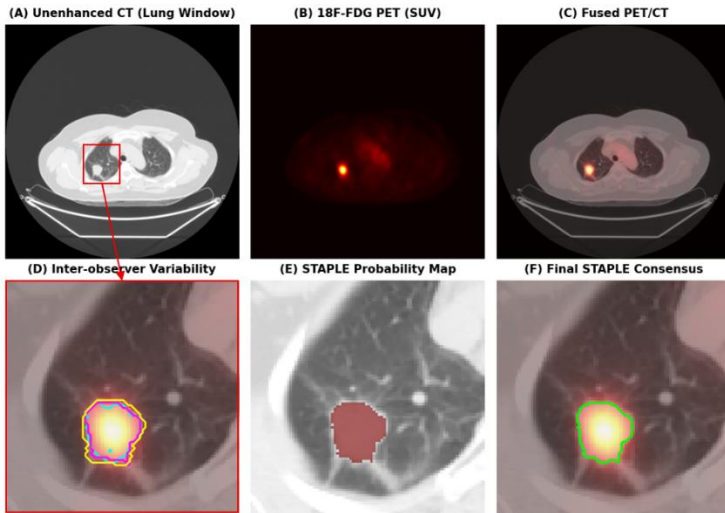
2.3. Phát triển nền tảng tự động phân đoạn ảnh PET/CT ứng dụng mô hình học sâu

Quy trình phân đoạn ung thư phổi không tế bào nhỏ (NSCLC) tự động được mô tả trong Hình 2.4. Quy trình đề xuất vận hành theo ba bước liên tiếp. Đầu tiên, ở giai đoạn chuẩn bị dữ liệu, các ảnh 3D PET/CT thô được tiền xử lý, đồng thời thuật toán STAPLE được áp dụng để tổng hợp kết quả gán nhãn từ ba chuyên gia thành một ảnh chuẩn đồng thuận (consensus ground truth), sau đó dữ liệu được chuyển sang định dạng NiFTI. Tiếp theo, trong giai đoạn huấn luyện, tập dữ liệu được tăng cường (augmented) và đánh giá bằng phương pháp kiểm định chéo K-fold phân tầng (stratified 5-fold cross-validation). Mô hình nnU-Net v2 cải tiến được huấn luyện và so sánh với các mô hình nền tảng (nnU-Net v2, Swin UNETR và 3D U-Net) trong cùng một cấu hình đồng nhất (hàm mất mát hỗn hợp, kích thước vùng cắt patch size $128 \times 128 \times 128$, bộ tối ưu hóa AdamW). Cuối cùng, ở giai đoạn suy luận và đánh giá, các kết quả dự đoán được tạo ra thông qua chiến lược cửa sổ trượt (sliding-window) kết hợp tăng cường dữ liệu tại thời điểm kiểm thử (test-time augmentation), kể đến là bước hậu xử lý để thu được mặt nạ dự đoán 3D (3D mask) cuối cùng. Hiệu năng của mô hình sau đó được đánh giá dựa trên các chỉ số không gian (DSC, HD95, ASSD) và độ tương quan lâm sàng thông qua thể tích khối u chuyển hóa (MTV). Hình ảnh PET và CT được tái lấy mẫu (resampling) về kích thước voxel đẳng hướng $2.0 \times 2.0 \times 2.0 \text{ mm}^3$; trong đó, phép nội suy spline bậc ba (cubic spline interpolation) được áp dụng cho dữ liệu cường độ pixel, còn phép nội suy láng giềng gần nhất (nearest-neighbor interpolation) được dùng cho các nhãn

mặt nạ (mask labels) nhằm bảo toàn tính nguyên vẹn của ranh giới khối u. Các kỹ thuật tăng cường dữ liệu (data augmentation) được áp dụng bao gồm xoay ngẫu nhiên, thay đổi tỷ lệ (scaling), và đặc biệt là bổ sung nhiễu Gaussian vào dữ liệu PET để mô phỏng đặc tính nhiễu Poisson vốn có trong ảnh y học hạt nhân. [Hình 2.4](#) minh họa dữ liệu hình ảnh đa phương thức đầu vào cùng quy trình khởi tạo nhãn chuẩn (ground-truth) nghiêm ngặt được áp dụng trong nghiên cứu này.



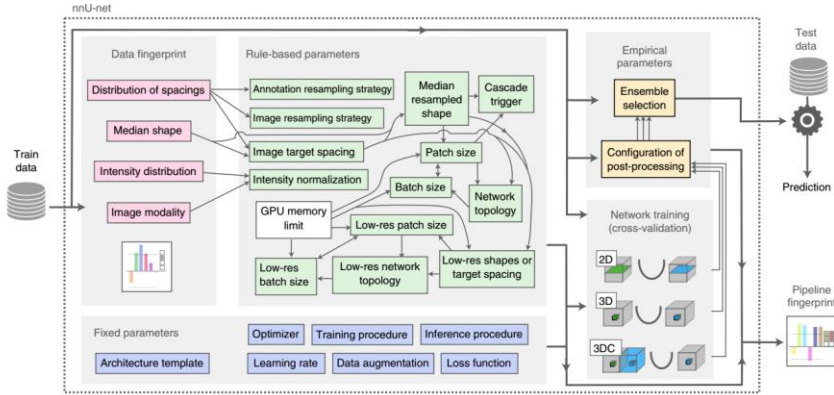
Hình 2.4. Nền tảng phân đoạn tự động



Hình 2.4: **Trực quan hóa dữ liệu PET/CT đa phương thức và quy trình khởi tạo nhãn chuẩn.** (A) Ảnh CT không cản quang; (B) Ảnh PET ^{18}F -FDG (bản đồ SUV); (C) Ảnh dung hợp PET/CT; (D) Sự khác biệt giữa ba chuyên gia khoanh vùng; (E) Bản đồ xác suất từ thuật toán STAPLE; và (F) Kết quả đồng thuận xác suất cuối cùng được sử dụng làm tiêu chuẩn tham chiếu.

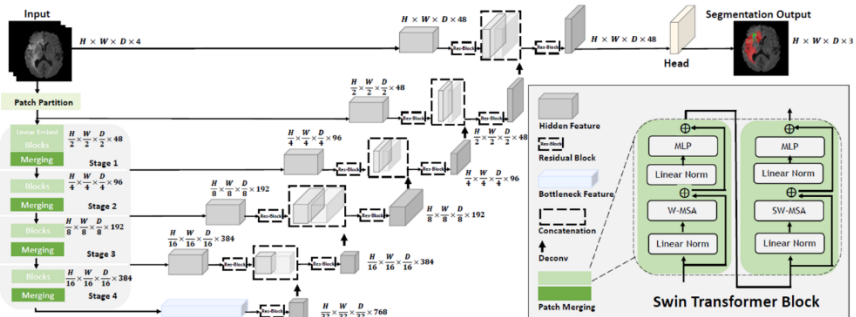
Ba kiến trúc học sâu chiến lược đã được lựa chọn để hiện thực hóa mục tiêu này:

- **nnU-Net v2 (Mô hình cơ sở):** Nhờ thể mạnh tự cấu hình thông minh theo đặc trưng dữ liệu mà không cần can thiệp thủ công, mô hình tiếp nhận đồng thời dữ liệu đa kênh PET và CT (kích thước khối đầu vào $128 \times 128 \times 128$ voxel), mang lại độ ổn định vượt trội trong việc xác định ranh giới giải phẫu.



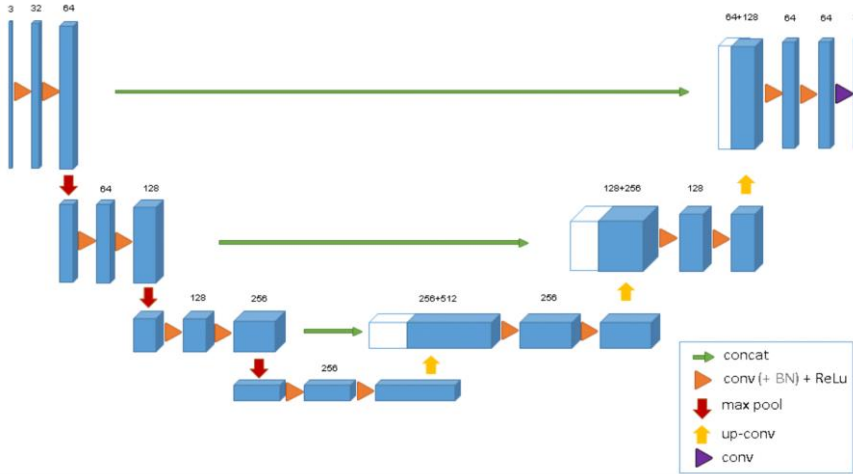
Hình 2.5: Cấu trúc nnU-Net

- **Swin UNETR (Kiến trúc đột phá ngữ cảnh):** Khắc phục triệt để giới hạn về trường thụ cảm cục bộ của CNN truyền thống bằng cách tích hợp cơ chế tự chú ý (self-attention) của Transformer. Qua việc phân rã dữ liệu 3D thành các token khối nhỏ và xử lý bằng các cửa sổ $7 \times 7 \times 7$, mô hình thiết lập thành công các mối phụ thuộc tầm xa, đặc biệt hiệu quả trong việc nhận diện các khối u NSCLC kích thước lớn hoặc có hình thái xâm lấn phức tạp, từ đó tối ưu hóa độ nhạy tổng thể.



Hình 2.6: Cấu trúc Swin UNETR

- **3D U-Net với ResNet backbone (Mô hình kiểm chứng):** Được triển khai trên nền tảng MONAI nhằm đánh giá hiệu quả dung hợp đặc trưng ở cấp độ bộ mã hóa. Sự chênh lệch hiệu năng giữa kiến trúc tiêu chuẩn này và phiên bản nnU-Net tinh chỉnh (đạt chỉ số Dice bút phá 83,44%) là minh chứng rõ nét cho thấy: nếu không qua học chuyển giao và tinh chỉnh sâu trên miền dữ liệu mục tiêu, các mô hình phổ thông rất khó vượt qua được rào cản dịch chuyển miền dữ liệu (domain shift).



Hình 2.7: Cấu trúc 3D U-Net

Toàn bộ quy trình được vận hành trên máy chủ trang bị card NVIDIA RTX Pro 6000 Blackwell Max-Q (96GB GDDR7), giúp duy trì kích thước vùng ảnh đầu vào đạt $128 \times 128 \times 192$ voxel ở độ phân giải gốc. Về cấu hình huấn luyện, mô hình nnU-Net tuân thủ lịch trình mặc định (tối đa 1000 epochs) và được khởi tạo bằng trọng số tiền huấn luyện tối ưu từ giải pháp đoạt giải AutoPET II 2023 trên Zenodo. Trong khi đó, các kiến trúc triển khai qua khung làm việc MONAI (3D U-Net và Swin UNETR) được huấn luyện trong 300 epochs bằng thuật toán AdamW (tốc độ học ban đầu 1×10^{-4} kết hợp suy giảm hàm cos, hệ số suy giảm trọng số 1×10^{-5} , với kích thước lô tổng thể bằng 8 nhờ cơ chế tích lũy gradient. Hàm mất mát hỗn hợp (hybrid loss function) đã được sử dụng để tối ưu hóa mạng thần kinh, kết hợp tính ổn định của hàm mất mát phân phối entropy chéo (cross-entropy loss) với khả năng xử lý hiện tượng mất cân bằng lớp (class imbalance) vốn là ưu thế của hàm mất mát Soft-Dice. Công thức được định nghĩa cụ thể như sau:

$$L = L_{CE} + L_{Dice} \quad (2.1)$$

Trong đó, các thành phần được tính toán theo công thức sau:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N y_i \log(p_i) \quad (2.2)$$

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N y_i p_i + \varepsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N p_i + \varepsilon} \quad (2.3)$$

Trong đó, y_i là nhãn thực tế (ground truth) tại voxel thứ i ; p_i là xác suất dự báo của mô hình tại voxel đó; N là tổng số lượng voxel trong không gian ảnh 3D; và ε là một hằng số làm mịn (smoothing constant) cực nhỏ nhằm ngăn chặn lỗi chia cho số không (0).

Quy trình đánh giá độ chính xác không gian và hiệu năng phát hiện tổn thương thần kinh nghiêm ngặt theo các hướng dẫn nghiêm ngặt từ bộ khung tiêu chuẩn Metrics Reloaded, với các phép toán được thực hiện thông qua thư viện `monai.metrics`:

- **DSC (Hệ số tương đồng Dice):** Đo lường mức độ chồng chập không gian giữa nhãn dự báo (P) và nhãn thực tế chuẩn (G).

$$DSC = \frac{2|P \cap G|}{|P| + |G|} \quad (2.4)$$

- **HD95 (Khoảng cách Hausdorff tại phân vị thứ 95):** Tính toán khoảng cách lớn nhất giữa hai tập hợp các điểm trên bề mặt nhưng loại trừ đi 5% số lượng điểm dị biệt (outliers) cao nhất, từ đó giảm thiểu tác động nhiễu gây ra bởi các vùng dự báo cô lập và có độ lệch lớn.

$$HD_{95} = \max \left(\max_{p \in P} \min_{g \in G} d(p, g), \max_{g \in G} \min_{p \in P} d(p, g) \right) \quad (2.5)$$

- **ASSD (Khoảng cách bề mặt đối xứng trung bình):** Đo lường khoảng cách trung bình từ bề mặt của vùng phân đoạn tự động đến bề mặt của vùng dán nhãn thủ công và ngược lại. Chỉ số ASSD càng thấp chứng tỏ đường biên dự báo càng bám sát vào nhãn gốc.

$$ASSD = \frac{\sum_{p \in P} d(p, G) + \sum_{g \in G} d(g, P)}{|P| + |G|} \quad (2.6)$$

CHƯƠNG 3. KẾT QUẢ VÀ BÀN LUẬN

3.1. Đánh giá bộ dữ liệu người Việt

Nghiên cứu sinh đã chọn lọc 416 bộ dữ liệu ảnh PET/CT của 416 bệnh nhân mắc NSCLC từ Bệnh viện K, Bệnh viện Ung bướu Hà Nội và Bệnh viện Ung bướu TP. Hồ Chí Minh. Để đảm bảo tính nhất quán và độ tin cậy trong ứng dụng lâm sàng, dữ liệu hình ảnh của mỗi bệnh nhân được phân đoạn độc lập bởi ba bác sĩ chuyên khoa ung thư. Chất lượng và mức độ đồng thuận của các nhãn phân

đoạn sau đó được đánh giá định lượng thông qua hai chỉ số tiêu chuẩn trong y tế là độ sai lệch thể tích khối u (VE) và chỉ số tương đồng Dice (DSC).

Độ sai lệch tương đối về thể tích khối u (VE) và Chỉ số tương đồng (DSC)

Độ sai lệch tương đối về thể tích khối u chính là sự khác nhau giữa thể tích khối u giữa hai dữ liệu phân đoạn của hai bác sỹ khác nhau. Thông qua việc tính thể tích của khối u từ tệp segmentation, nghiên cứu sinh đã tính toán giá trị độ lệch về thể tích như trong **Bảng 3.1**. với kết quả độ lệch trung bình khoảng 0.2. Mức độ sai lệch thấp này — đặc biệt là giữa bác sĩ thứ nhất và bác sĩ thứ hai — cho thấy sự đồng thuận cao giữa các người quan sát, tương đương hoặc thậm chí vượt trội so với kết quả từ các nghiên cứu PET/CT tương tự.

Bảng 3.1. So sánh độ lệch về thể tích của tập hợp phân đoạn giữa các bác sỹ

So sánh giữa bác sỹ	Khoảng giá trị VE	Số lượng (bệnh nhân)	Tỷ lệ (%)
Bác sỹ 1 và Bác sỹ 2	0,0-0,1	30	10,0
	0,1-0,2	179	59,6
	0,2-0,3	91	30,4
Bác sỹ 2 và Bác sỹ 3	0,0-0,1	21	7,00
	0,1-0,2	148	49,3
	0,2-0,3	131	43,7
Bác sỹ 1 và Bác sỹ 3	0,0-0,1	29	9,7
	0,1-0,2	157	52,3
	0,2-0,3	114	38,0

Chỉ số tương đồng (Dice Similarity Coefficient-DSC) là chỉ số phổ biến nhất để đánh giá mức độ trùng khớp giữa hai tập hợp phân đoạn, chẳng hạn như giữa phân đoạn của hai chuyên gia hoặc giữa chuyên gia và mô hình.

$$DSC = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{3.1}$$

Trong đó A và B là hai tập hợp thể tích được phân đoạn, $\cap$ là phần giao nhau giữa chúng, và $|A|$, $|B|$ là kích thước của từng tập hợp. Để đánh giá chất lượng của bộ dữ liệu, giá trị của DSC dao động từ 0 (không trùng khớp) đến 1 (trùng khớp hoàn toàn). Đối với phân đoạn khối u, các chuyên gia thường yêu cầu DSC đạt ít nhất **0,7 đến 0,8** để coi là đủ trùng khớp (**Bảng 3.2**).

Thông qua tính toán, nghiên cứu sinh đã xác định 100% dữ liệu phân đoạn của các bệnh nhân đều đạt độ trùng khớp từ 70% trở lên, vượt qua ngưỡng chấp nhận theo tiêu chuẩn của Image Biomarker Standardization Initiative với $DSC \geq 0,7$ được coi là dấu hiệu của sự đồng thuận tốt giữa các chuyên gia trong phân đoạn y học, đặc biệt với các khối u phức tạp như ung thư phổi không tế bào nhỏ (NSCLC). Về tương quan giữa các bác sĩ, cặp bác sĩ 1 và bác sĩ 2 đạt độ trùng khớp tốt nhất là 82,3%, cao hơn so với mức trung bình 0,75-0,80 trong nghiên cứu của Lambin cho thấy sự phối hợp hiệu quả, có thể do kinh nghiệm tương đồng hoặc sử dụng công cụ như 3D Slicer.

Bảng 3.2. So sánh chỉ số tương đồng của tập hợp phân đoạn giữa các bác sỹ

So sánh giữa bác sỹ	Khoảng DSC	Số lượng (bệnh nhân)	Tỷ lệ (%)
Bác sỹ 1 và Bác sỹ 2	0,9-1,0	32	10,67
	0,8-0,9	183	61,00
	0,7-0,8	85	28,33
Trung bình	0,82±0,08		
Bác sỹ 2 và Bác sỹ 3	0,9-1,0	25	8,33
	0,8-0,9	154	51,33
	0,7-0,8	121	40,33
Trung bình	0,78±0,06		
Bác sỹ 1 và Bác sỹ 3	0,9-1,0	34	11,33
	0,8-0,9	169	56,33
	0,7-0,8	97	32,33
Trung bình	0,80±0,07		
Trung bình bộ dữ liệu	0,80±0,07		

3.2. Đánh giá công cụ đọc và tiền xử lý ảnh PET/CT

Giao diện hiển thị chính của chương trình được chia thành bốn khu vực (Hình 3.1) để lần lượt hiển thị lát cắt ngang, lát cắt đứng và lát cắt bên của dữ liệu.

PHẦN MỀM ĐỌC DỮ LIỆU HÌNH ẢNH DICOM

File | Chỉnh sửa | Xem | Thông tin

THÔNG TIN BỆNH NHÂN

Tên bệnh nhân: BUI THỊ NHỊEM 1930-C34

Tuổi: 092Y

ID: 22025300

Giới tính: F

Bác sỹ: 22025300

THÔNG TIN ẢNH KHỐI U

Liều chụp PET: mCi

Kích thước khối u: mm3

Tọa độ khối u:

SUV cực đại:

Thời gian chụp: s

PHÂN ĐOẠN HÌNH ẢNH | HIỂN THỊ KHỐI U

Hình 3.1. Giao diện chương trình và hiển thị ảnh DICOM

Một phần hiển thị còn lại của phần mềm sẽ được dùng để hiển thị hình ảnh khối u đã được bác sĩ phân đoạn thủ công. Phần mềm cho phép có thể hiển thị nhiều lát cắt khác nhau trong bộ dữ liệu bằng cách di chuyển thanh dịch ở cửa sổ hiển thị tương ứng. Phần mềm có thể cài đặt được trên các máy tính sử dụng hệ điều hành Windows hoặc Linux, nhìn chung, phần mềm hiển thị đã đáp ứng được đầy đủ các yêu cầu đề ra của luận án. Tốc độ xử lý của phần mềm còn chưa được tốt vì được xây dựng trên nền tảng đơn nhân, thời gian xử lý các bộ dữ liệu lớn sẽ rất lâu. Phần mềm còn cần được tiếp tục phát triển khả năng xử lý đa luồng, xử lý dữ liệu kết hợp với card đồ họa để cho thời gian xử lý thực tiễn hơn. Trên cơ sở phần mềm đọc dữ liệu, các mô đun học máy có thể được tích hợp thêm vào để tạo nên phần mềm hỗ trợ chẩn đoán hoàn chỉnh.

3.3. Đánh giá chất lượng nền tảng tự động phân đoạn hình ảnh ứng dụng mô hình học sâu

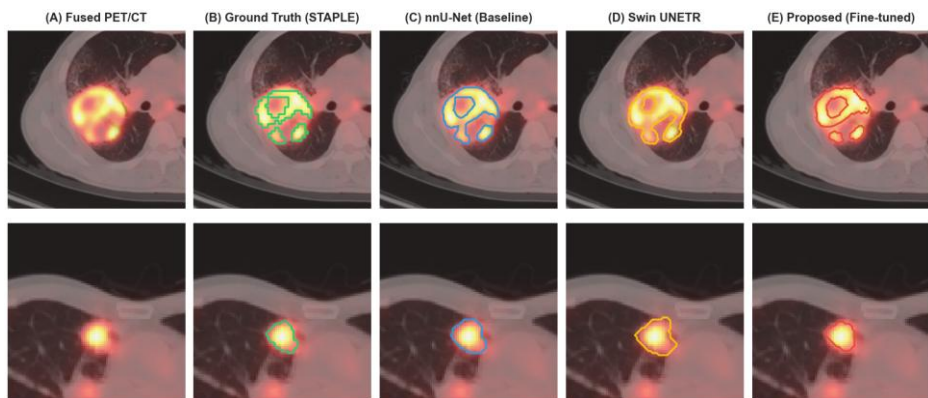
Có thể thấy rõ từ [Bảng 3.4](#), mô hình nnU-Net (tiền huấn luyện + tinh chỉnh) đã chứng minh được sự vượt trội rõ rệt. Việc tận dụng các đặc trưng tổng thể kết hợp cùng quy trình tinh chỉnh trên nguồn dữ liệu bản địa đã giúp mô hình này đạt được chỉ số DSC cao nhất, ở mức $83,44 \pm 6,5\%$. Đồng thời, mô hình này cũng ghi nhận các sai số khoảng cách không gian thấp nhất, với chỉ số HD95 chỉ ở mức $5,1 \pm 3,6$ mm và ASSD đạt $2,0 \pm 2,2$ mm, thể hiện khả năng bám sát cực tốt và chính xác vào các đường biên của khối u. Độ chính xác (precision) của kiến trúc tinh chỉnh này cũng đạt hiệu năng cao nhất ($89,60 \pm 8,2\%$), cho thấy khả năng kiểm soát xuất sắc đối với các trường hợp dương tính giả tại các vùng hấp thu FDG sinh lý.

Bảng 3.3: So sánh định lượng hiệu năng phân đoạn giữa các kiến trúc học sâu

Model	DSC (%) ↑	HD95 (mm) ↓	ASSD (mm) ↓	Sensitivity (%) ↑	Precision (%) ↑	p-value (vs. Fine-tuned)
3D U-Net (MONAI)	65.0 ± 10.2	15.8 ± 8.5	6.2 ± 2.1	66.2 ± 10.5	68.1 ± 12.0	< 0.001
Swin UNETR	71.2 ± 8.6	9.5 ± 6.2	3.1 ± 1.5	80.4 ± 8.1	77.5 ± 10.2	< 0.001
nnU-Net v2	76.6 ± 7.5	7.3 ± 8.6	2.5 ± 2.3	77.2 ± 7.4	81.3 ± 8.5	< 0.01
nnU-NET (pretrained + fine-tuned)	83.44 ± 6.5	5.1 ± 3.6	2.0 ± 2.2	79.19 ± 8.2	89.60 ± 8.2	-

Lưu ý: Ký hiệu ↑ biểu thị giá trị càng cao càng tốt; ↓ biểu thị giá trị càng thấp càng tốt.

[Hình 3.2](#) trình bày các kết quả định tính mang tính đại diện từ hai ca bệnh thực tế, so sánh một cách trực quan kết quả đầu ra phân đoạn của các kiến trúc huấn luyện khác nhau.



Hình 3.2. Kết quả định tính tiêu biểu với các kiến trúc huấn luyện khác nhau

Qua quan sát trực quan, các mô hình nnU-Net cơ sở và Swin UNETR gặp nhiều khó khăn trong việc phác họa một cách chính xác các đường biên khối u phức tạp. Ngược lại, mô hình nnU-Net tinh chỉnh thể hiện độ chồng chấp tốt nhất đối với kết quả đồng thuận STAPLE (STAPLE consensus), đặc biệt là tại các vùng có mức hấp thu FDG không đồng nhất (heterogeneous). Nhìn chung, với độ chính xác không gian và độ chính xác phân loại (precision) vượt trội, mô hình nnU-Net tinh chỉnh đã được lựa chọn để sử dụng cho toàn bộ các phân tích chuyên sâu tiếp theo (downstream analyses) trong nghiên cứu này.

Khi tiến hành so sánh độc lập với nhân chú giải của từng bác sĩ (được trình bày ở Bảng 3.4), mô hình nnU-Net tinh chỉnh đã đạt được chỉ số Dice trung bình vô cùng ấn tượng là 0,818. Cụ thể, mô hình đạt giá trị DSC lần lượt là 0,815, 0,822 và 0,818 khi đánh giá đối chiếu với Người đánh giá 1 (Rater 1), Người đánh giá 2 (Rater 2) và Người đánh giá 3 (Rater 3); đồng thời đạt đỉnh với giá trị DSC là 0,834 khi so sánh với nhân đồng thuận STAPLE.

Bảng 3.4. Phân tích DSC về sự phù hợp giữa mô hình

Model	Rater 1	Rater 2	Rater 3	STAPLE
nnU-Net (pretrained + fine-tuned)	0.815	0.822	0.818	0.834

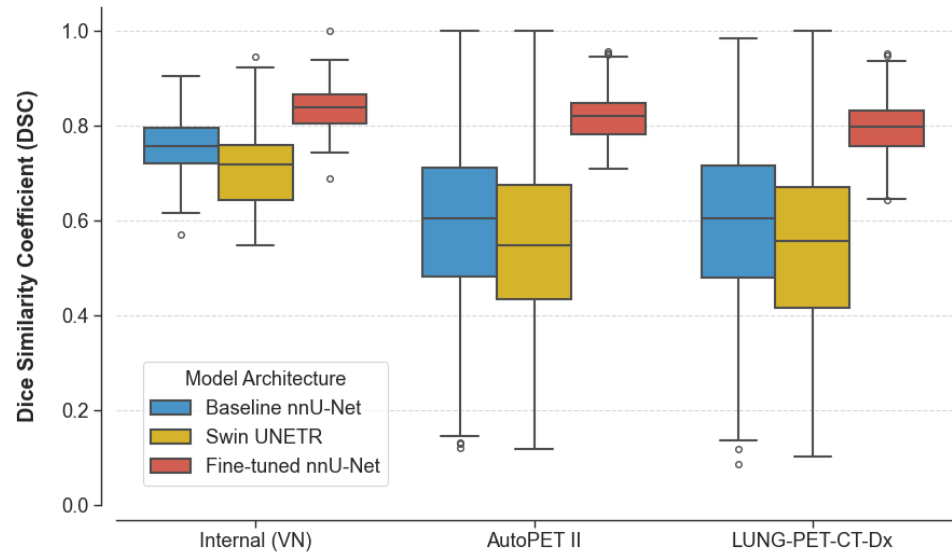
Như được minh họa ở Bảng 3.5, Với mức sụt giảm hiệu năng trung bình cực kỳ thấp — chỉ vón vện -3,5% — khung làm việc dựa trên chiến lược tinh chỉnh đã chứng minh được năng lực tổng quát hóa vượt xa các kiến trúc chỉ được huấn luyện cục bộ từ đầu.

Sự vượt trội về khả năng tổng quát hóa của mô hình tinh chỉnh được minh chứng rõ nét hơn thông qua phân phối thống kê ở Hình 3.3. Biểu đồ chỉ ra rằng khi đối mặt với hiện tượng dịch chuyển miền trên các bộ dữ liệu độc lập bên ngoài, các kiến trúc cơ sở (nnU-Net baseline và Swin UNETR) không chỉ bị sụt giảm nghiêm trọng về giá trị trung vị (median values) mà còn bộc lộ mức độ phân tán dữ liệu khổng lồ. Khoảng tứ phân vị (IQR) của các mô hình này bị giãn rộng đáng kể, đi kèm với hàng loạt điểm dữ liệu dị biệt (outliers) có hiệu năng lao dốc xuống các

ngưỡng cực thấp ($< 0,4$). Ngược lại, mô hình nnU-Net tinh chỉnh vẫn duy trì được độ ổn định đáng kinh ngạc. Biên độ phân tán của mô hình đề xuất trên cả hai bộ dữ liệu AutoPET II và LUNG-PET-CT-Dx đều rất hẹp và tương đương với mức phân phối ghi nhận trên bộ dữ liệu nội bộ.

Bảng 3.5. Hiệu suất tổng quát hóa không cần huấn luyện của các mô hình

Model	Internal (VN)	AutoPET II	LUNG-PET-CT-Dx	Average Drop
nnU-Net v2 (baseline)	0.766	0.612	0.598	-21.0%
Swin UNETR	0.712	0.577	0.560	-20.1%
nnU-Net (pretrained + fine-tuned)	0.834	0.811	0.798	-3.5%



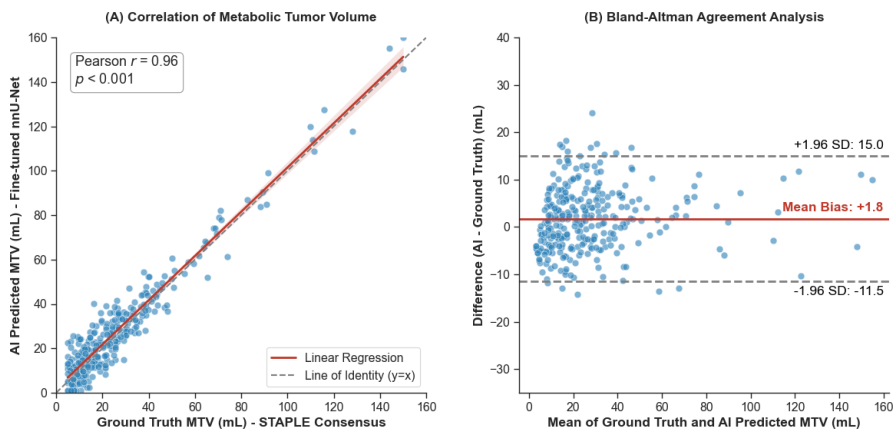
Hình 3.3. Phân bố DSC trên các tập dữ liệu nội bộ và bên ngoài.

Bảng 3.6 cho thấy những lợi ích sâu sắc mà chiến lược học chuyển giao (transfer learning) mang lại trên cả hai phương diện: hiệu năng mô hình và thời gian tính toán. Việc huấn luyện mô hình nnU-Net cơ sở từ đầu (train from scratch) cho chỉ số DSC đạt 0,766 và khoảng cách đường biên HD95 đạt 7,30 mm, đòi hỏi thời gian xử lý xấp xỉ 16 giờ cho mỗi tập kiểm định. Trong khi đó, quy trình tinh chỉnh (fine-tuning) đã thúc đẩy đáng kể độ chính xác phân đoạn, nâng chỉ số DSC lên mức 0,834, đồng thời rút ngắn rõ rệt chỉ số HD95 xuống còn 5,10 mm.

Bảng 3.6. Kết quả nghiên cứu loại bỏ đánh giá tác động của học chuyển giao lên độ chính xác phân đoạn và hiệu quả huấn luyện.

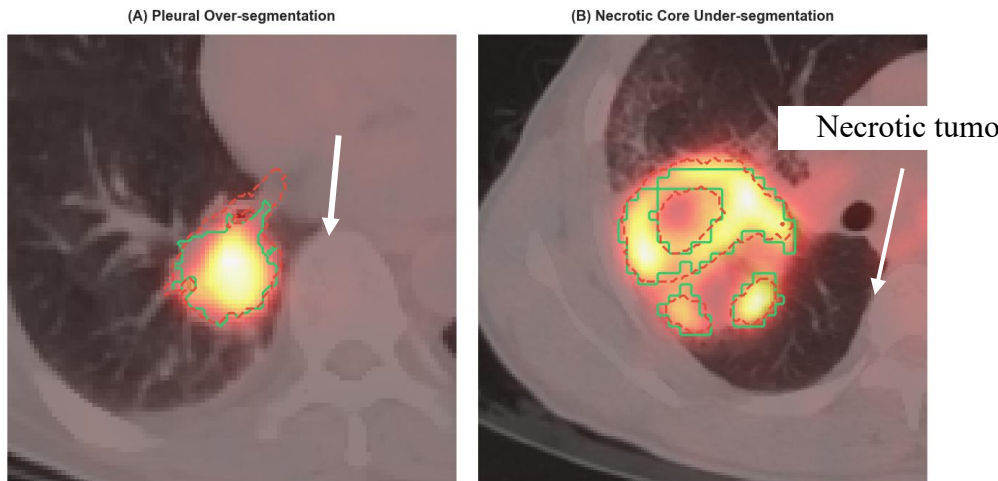
Configuration	DSC	HD95 (mm)	Training time (hours/fold)
Baseline nnU-Net	0.766	7.30	16
Fine-tuned nnU-Net	0.834	5.10	10

Hình 3.4a trình bày các kết quả phân tích thống kê cho thấy một mối tương quan tuyến tính cực kỳ mạnh mẽ giữa kết quả dự báo của AI và kết quả đánh giá thủ công của bác sĩ, đạt hệ số tương quan Pearson lên tới $r = 0,96$ ($p < 0,001$). Bên cạnh đó, biểu đồ phân tích Bland-Altman (Hình 3.4b) minh chứng cho mức độ đồng thuận lâm sàng ở mức cao, ghi nhận độ sai lệch trung bình (mean bias) chỉ ở mức +1,8 mL. Giới hạn đồng thuận 95% (95% limits of agreement) nằm trong một khoảng hẹp từ -11,5 mL đến +15,0 mL, đây là mức sai số hoàn toàn chấp nhận được trong phạm vi dung sai cho phép của quy trình lập kế hoạch xạ trị thực tế.



Hình 3.4: Phân tích sự tương đồng lâm sàng của MTV được trích xuất bởi mô hình nnU-Net tinh chỉnh so với sự đồng thuận của chuyên gia STAPLE.

Hình 3.5 minh họa sự khác biệt giữa nhãn đồng thuận STAPLE (nét liền xanh lá) và mô hình nnU-Net tinh chỉnh (nét đứt đỏ), với các mũi tên trắng chỉ ra vùng bất đồng nhất. Dù có độ bền vững tổng thể cao, mô hình vẫn bộc lộ hạn chế tại các ranh giới phức tạp do hiệu ứng thể tích từng phần làm mờ đường biên, dẫn đến phân đoạn quá mức ở vùng tiếp giáp khối u và màng phổi (Hình 3.5A). Ngược lại, Hình 3.5B cho thấy hiện tượng phân đoạn thiếu hụt đối với khối u có lõi hoại tử không hấp thu FDG, khi mô hình chỉ khu trú tại viền tăng chuyển hóa mà bỏ sót thể tích hoại tử bên trong đã được các chuyên gia phác họa. Những trường hợp này làm nổi bật thách thức trong việc phân đoạn các khối u có tính đa dạng cao.



Hình 0.5. Phân tích định tính sự không nhất quán phân đoạn ở các vùng ranh giới giải phẫu phức tạp: (A) phân đoạn quá mức màng phổi; (B) phân đoạn thiếu lõi khối u hoại tử không hấp thụ FDG

KẾT LUẬN

Luận án đã thực hiện được các mục tiêu và nội dung nghiên cứu đề ra, các kết quả nghiên cứu chính thể hiện đóng góp mới bao gồm:

- 1) Luận án đã xây dựng bộ dữ liệu PET/CT lồng ngực đa trung tâm quy mô lớn từ 416 bệnh nhân tại Bệnh viện K, Bệnh viện Ung Bướu Hà Nội và Bệnh viện Ung Bướu TP.HCM, bao gồm 300 ca ung thư phổi không tế bào nhỏ (NSCLC) và 116 ca đối chứng. Mỗi ca bệnh chứa từ 200–300 lát cắt CT/PET cùng đầy đủ thông tin lâm sàng (tuổi, giới tính, cân nặng, liều ^{18}F -FDG và giai đoạn bệnh). Đặc biệt, toàn bộ 300 ca NSCLC đều được chuẩn hóa nghiêm ngặt, bao gồm ảnh gốc đi kèm 03 kết quả khoanh vùng tổn thương độc lập từ 03 bác sĩ chuyên khoa, tạo cơ sở dữ liệu có độ tin cậy cao phục vụ huấn luyện mô hình.
- 2) Luận án đã xây dựng thành công chương trình tự động đọc, xử lý và xác định biên dạng khối u từ dữ liệu phân đoạn thủ công của bác sĩ, phục vụ lưu trữ và huấn luyện mô hình học máy. Kết quả trích xuất bao gồm cả thông tin tổng quát của khối u và chi tiết trên từng lát cắt.
- 3) Luận án đã phát triển và kiểm định thành công hệ thống tự động phân đoạn khối u ung thư phổi không tế bào nhỏ (NSCLC) dựa trên mô hình nnU-Net v2 tinh chỉnh. Được tối ưu hóa trên bộ dữ liệu 416 ca bệnh tại Việt Nam,

mô hình đạt hiệu năng vượt trội so với các kiến trúc Swin UNETR và nnU-Net cơ sở, thể hiện tính nhất quán lâm sàng cao (chỉ số DSC đạt 0,834 so với nhãn đồng thuận STAPLE) và giảm tỷ lệ dương tính giả vùng hấp thu sinh lý xuống còn 6%. Đặc biệt, nhờ chiến lược học chuyển giao, mô hình duy trì hiệu năng ổn định trên các tập dữ liệu độc lập bên ngoài (AutoPET II và LUNG-PET-CT-Dx) với mức sụt giảm trung bình chỉ -3,5%, đồng thời tiết kiệm 37,5% thời gian huấn luyện. Kết quả này khẳng định việc bản địa hóa các mô hình học sâu tiên tiến là giải pháp tối ưu để vượt qua rào cản dịch chuyển miền, cung cấp một công cụ tự động hóa đáng tin cậy phục vụ chuẩn hóa quy trình phân chia giai đoạn bệnh và lập kế hoạch xạ trị.

DANH SÁCH BÀI BÁO KHOA HỌC

Danh sách các bài báo quốc tế

1. **Ho Quang Tuan**, Tran Thuy Duong, Bui Ngoc Ha, Nguyen Huu Quyet, Le Van Tinh, Chu Van Tuynh, Vo Khac Nam, Le Thi Quynh Dao, Chu Van Luong, Dang Thi My Linh, Duong Thi Nhung, Nguyen Dang Nguyen, Vo Quynh Trang; *Development of a Vietnamese PET/CT Dataset for Machine Learning-Based Analysis of Non-Small Cell Lung Cancer Images*; Atom Indonesia, Vol. 51 No.2 (2025), 137-143. (<https://atomindonesia.brin.go.id/index.php/aij/article/view/1645>)
2. Bui Tien Hung, Tran Thuy Duong, Bui Ngoc Ha, **Ho Quang Tuan**; *A Monte Carlo simulation of Gamma ART-6000 devices to investigate beam characteristics in a homogeneous environment using Geant4 simulation toolkits*; Radiation Physics and Chemistry - 2025, Vol. 229 (<https://doi.org/10.1016/j.radphyschem.2025.112542>)
3. Ioannis Delakis, Zoe Brady, Md Nahid Hossain, Rath Beauta, Sann Chanthoubopha, Chhoeurt Koeurn, Yun Seang Yi, Rajesh Kumar, Lukmanda Evan Lubis, Supriyanto Ardjo Pawiro, Indah Lestariningsih, Hideyuki Mizuno, Kosuke Matsubara, Keiichi Akahane, Siti Normasitah Masduki, Siti Norsyafiqah Mohd Mustafa, Kanchan Prasad Adhikari, Darin O’Keeffe, Atif Masood, Luwis Gabriel, Sanja Ipu, Melanie Dela Fuente Marquez, Ramon Carlo L. Cruzpero, Soma Somanesan, Anchali Krisanachinda, **Tuan Quang Ho**, Olivera Ciraj-Bjelac; *Status of the medical physics profession in diagnostic radiology and image-guided interventional procedures in the Asia–Pacific region: initial findings from*

Danh sách các bài báo trong nước

1. **Hồ Quang Tuấn**, Đặng Thị Mỹ Linh, Dương Thị Nhung, Nguyễn Hữu Quyết, Nguyễn Ngọc Quỳnh, Bùi Tiến Hưng, Lê Xuân Chung, Trần Thuỳ Dương, Bùi Ngọc Hà, Nguyễn Văn Thái, Phạm Lâm Sơn, Lê Văn Tình, Chu Văn Lương, Chu Văn Tuynh, Bùi Xuân Cường, Hoàng Văn Toán, Võ Khắc Nam, Nguyễn Trung Hiếu, Đặng Thị Minh Tâm, Lê Thị Quỳnh Dao, Đặng Minh Phước; *Nghiên cứu thuật toán định vị và phân đoạn khối u phổi từ các thể tích lồng ngực PET/CT dựa trên phân tích đặc điểm hình ảnh*, Tạp chí Y học Việt Nam, 2023, tập 533 số Đặc biệt.
2. Bui Ngoc Ha, Nguyen Trung Kien, Tran Thi Thuy Duong, Tran Ngoc Toan, **Ho Quang Tuan**, Pham Cam Phuong, Pham Van Thai, Vu Thi Trang *Development of Brain MRI Image Segmentation program using UNET++ network for radiosurgery planning*; Communications in Physics, Journal of Nuclear Science and Technology, 2023, Vol. 13 No. 1. (<https://doi.org/10.53747/nst.v13i1.398>)